



Inventor: Samir K. Brahmachari
 Application No. 09/707,919
 Title: METHOD OF DETECTION OF
 ALLELIC VARIANTS OF SCA2 GENE
 Client/Matter No. 39562-175772

1/13

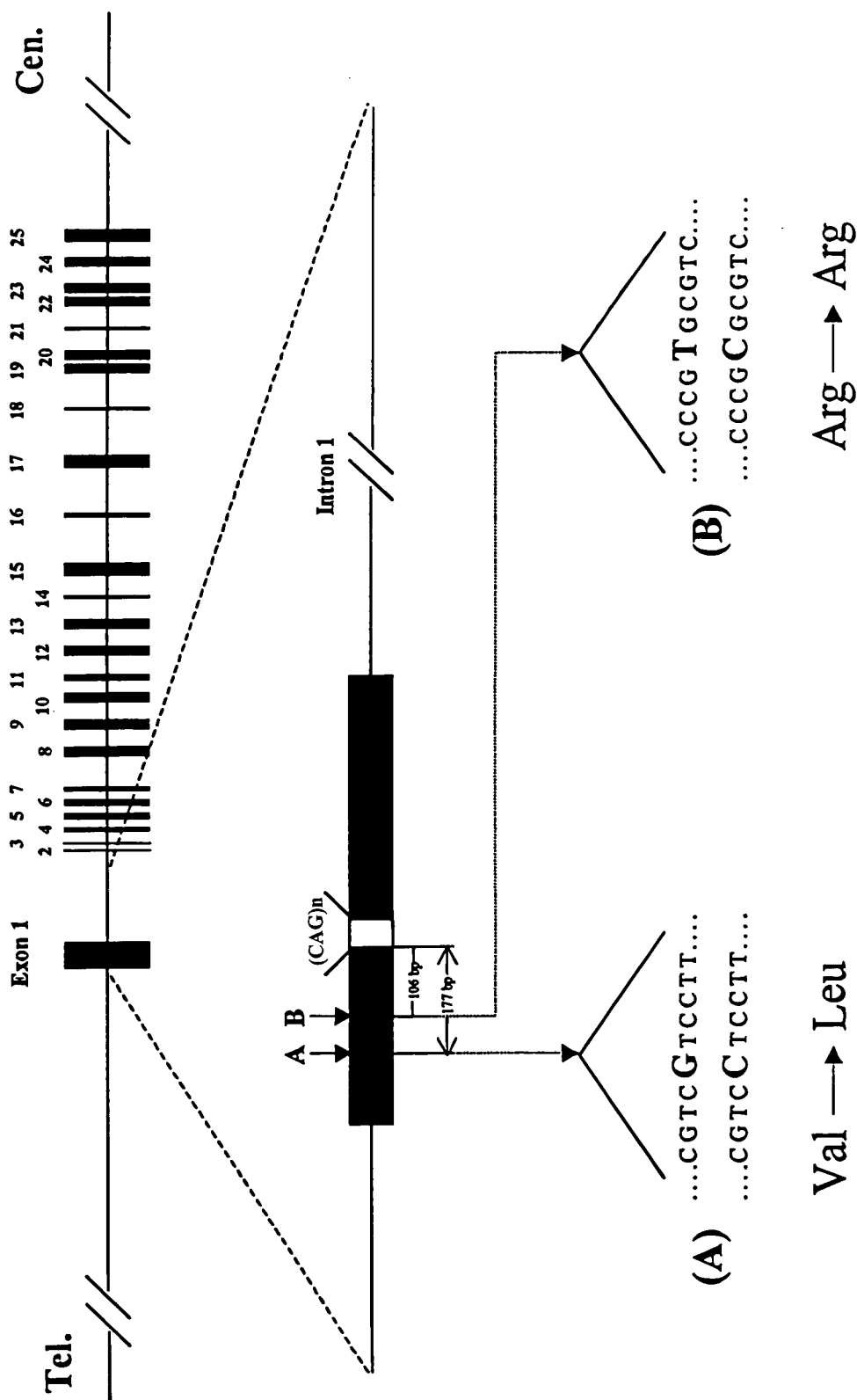
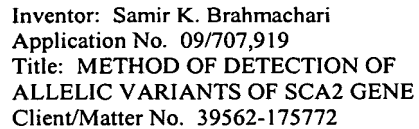


FIG 1



2/13

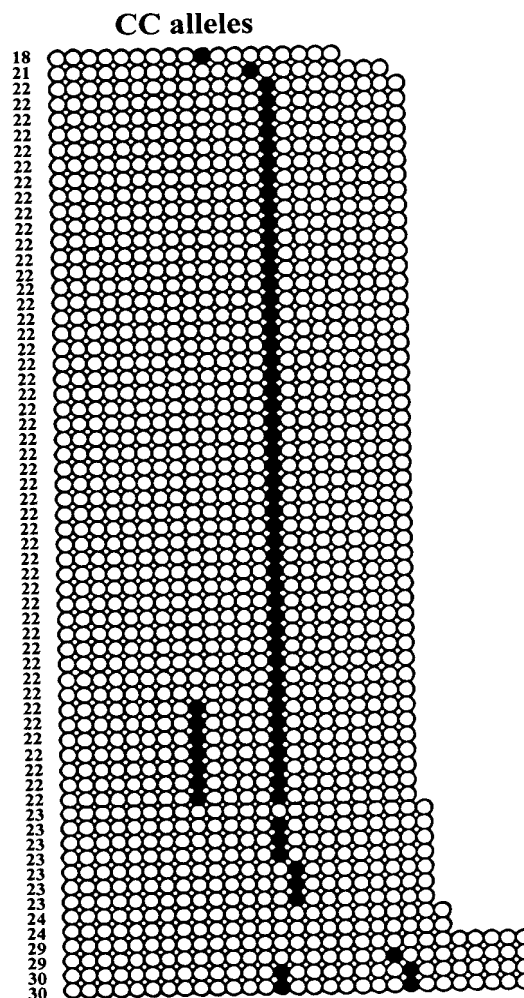
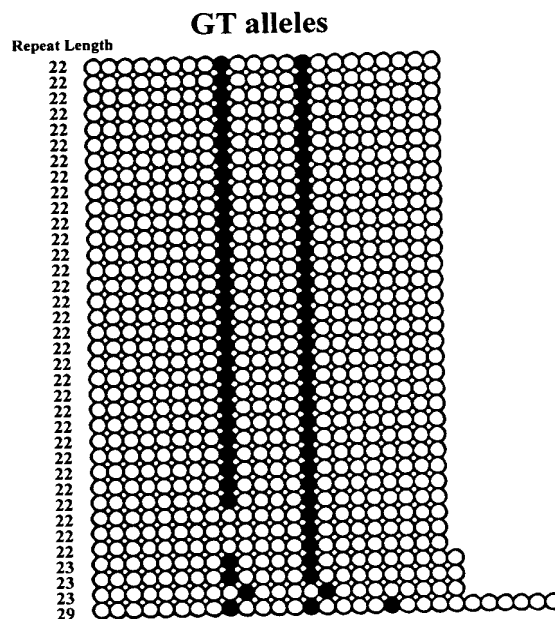


FIG 2



3/13

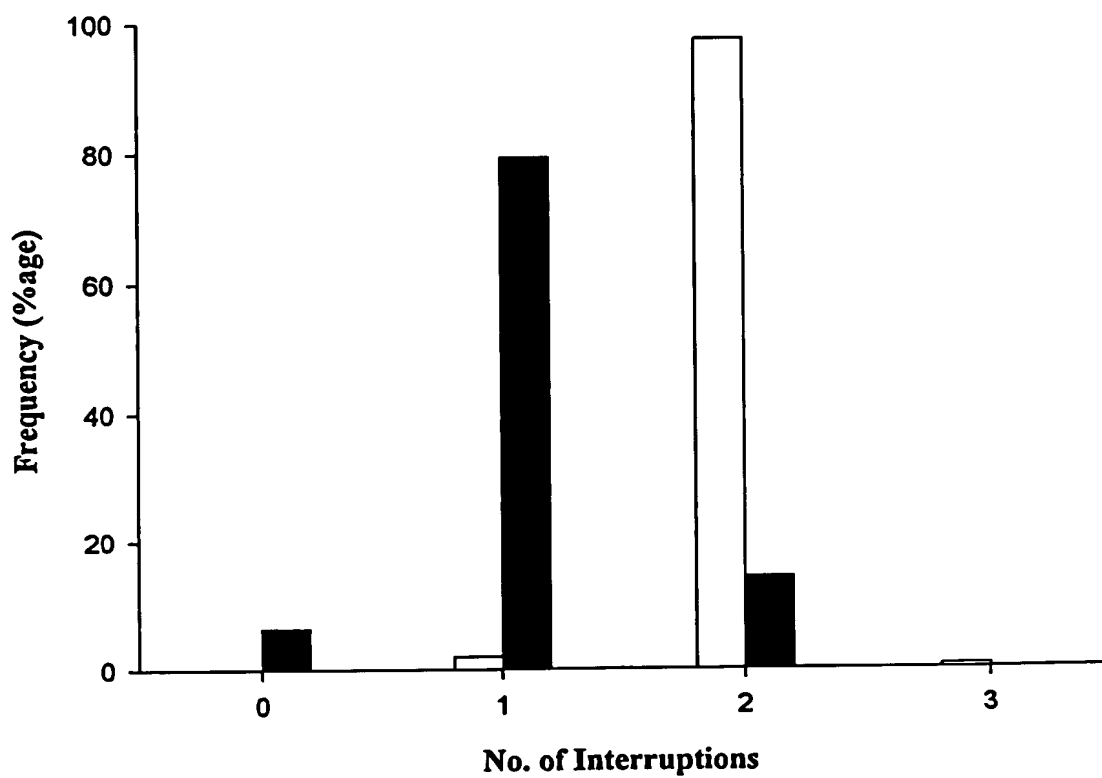


FIG. 3



4/13

SNP Details

SNP:

Handle|local_snp_id: FGU-CBT | SKB.2K.1.1
NCBI Assay Id(ss#): 869704
Reference SNP Id(rs#): 695871

STS Information: Not submitted

From SNP Database:

Submitter Handle: FGU-CBT
Submitter Batch ID: SKB.2K.1
Release Date: Aug. 2 2000 2:53PM
Molecular type: Genomic
No. of Chromosomes sampled: 215
Synonym defined:
Organism: Homo sapiens
Population: INDPOP
Submitter Method ID: SCA2-SNP
Citation:

Single Nucleotide Polymorphism in SCA2 Gene.

View citation details

NCBI Assay ID: 869704
Submitter SNP ID: SKB.2K.1.1
Synonyms:
LOCUSID: 6311
Submitter STS ID:
STS Accession: not available
GenBank Accession: U70323
Gene Name: Human ataxin-2 gene/ Spinocerebellar ataxia 2 (SCA2) gene
Length: 459

Flanking Sequence Information:

5'Assay: CTCCGCCTCA GACTGTTTTG GTAGCAACGG CAACGGCGGC GCGCGTTTC GGCCCGGCTC
CCGGCGGCTC CTTGGTCTCG GCGGGCCTCC CCGCCCCTTC GTCGTC (SEQ ID NO: 14)

Observed: G/C

3'Assay: TCCTTCTCCC CCTCGCCAGC CCGGGCGCCC CTCCGGCCGC GCCAACCCGC GCCTCCCCGC
TCGGCGCCCG TCGTCCCCG CCGCGTTCCG GCGTCTCCTT GGCGCGCGCT GCTCCCGGCT
GTCCCCGCC GCGGTGCGAG CCGGTGTATG GGCCCCCTAC CATGTCGCT (SEQ ID NO: 15)

3'Flank: GAAGCCCCAG CAGCAGCAGC AGCAGCAGCA GCAACAGCAG CAGCAGCAAC AGCAGCAGCA
GCAGCAGCAG CAGCCGCCGC CCGCGGCTGC CAATGTCCGC AAGCCCGGCG GCAGCGGCCT
TCTAGCGTCG CCCGCCGCC GCGCTTCGCC GTCCTCGTCC TCGGTCTCCT CGTCCTCGGC
CAC (SEQ ID NO: 16)

Allele Frequency Information:

POP Batch Id: SKB.2K.1:
Submitter Handle: FGU-CBT
Submitter Method ID: SCA2-SNP
Citation:
Single Nucleotide Polymorphism in SCA2 Gene.

Handle|PopulationID: FGU-CBT| INDPOP
No. of Chromosomes Sampled: 215

Allele: C = 0.293 / G = 0.707

FIG 4A



5/13

Reference SNP Record

NCBI SNP ID: rs695871

NCBI Resource Links

GenBank: [U70323](#)

LocusLink: no link established

Integrated Maps: *under construction*

Submitter records for this ID:

| Assay ID | Handle Local Submitter ID | Release Date |
|--------------------------|-----------------------------|-------------------|
| ss869704 | FGU-CBT SKB.2K.1.1 | Aug 2 2000 2:53PM |

Variation Summary:

Assay sample size (number of chromosomes) : 430
Population data sample size (number of chromosomes) :
Total number of populations with frequency data: 1
Total number of individuals with genotype data: 0
Average estimated heterozygosity: 0.414
Average Allele Frequency:

| | |
|----|-------|
| C: | 0.293 |
| G: | 0.707 |

Validation Summary:

Marker displays Mendelian segregation: UNKNOWN
PCR results confirmed in multiple reactions: YES
Homozygotes detected in individual genotype data: UNKNOWN
Insufficient genotype data to compute the goodness of fit to
Hardy-Weinberg
Insufficient data to compute individual x genotype consistency
measures
Validation status: *under construction*

FIG 4B



6/13

Submitter Contact Details

This batch's contact information:

handle: FGU-CBT
name: Shweta Choudhry
fax: +91-11-7257471
phone: +91-11-7416489
email: shwetachoudhry@hotmail.com
lab: Functional genomics Unit
institution: Centre for Biochemical Technology (CSIR)
address: Delhi University Campus, Mall Road, Delhi- 110007, India

Handle information for the lab:

handle: FGU-CBT
name: Prof. Samir K. Brahmachari
fax: +91-11-7257471
phone: +91-11-7416489
email: skb@cbt.res.in
lab: Functional genomics Unit
institution: Centre for Biochemical Technology (CSIR)
address: Delhi University Campus, Mall Road, Delhi- 110007, India

FIG. 4C

SNP Population Details - 558

Submitter Population Handle: FGU-CBT
Submitter population ID: INDPOP
Population Text:

Continent: Asia
Nation: India

[View SNP used on this population](#)

FIG 4D



7/13

SNP Method Details - 564

Submitter Method Handle: FGU-CBT
Submitter Method ID: SCA2-SNP
Method Text:

The region containing the SNP was PCR amplified using the primers SCA2-FP3 (5' CTCCGCTCAGACTGTTTTGGTAG 3') (SEQ ID NO: 1) and SCA2-RP3 (5' GTGGCCGAGGACGAGGAGAC 3') (SEQ ID NO. 2). Approximately 100ng of genomic DNA was amplified in a 50ml reaction volume containing a final concentration of 5mM Tris, 25mM KCl, 0.7mM MgCl₂, 0.05% gelatin, 20pmol of each primer and 0.5U of Taq DNA polymerase. Samples were denatured at 94oC for 3 min followed by 35 cycles of denaturation (94oC, 45sec), annealing (52oC, 30sec), extension (72oC, 45sec) and a final extension of 7 min at 72oC in a Perkin Elmer GeneAmp PCR System 9600. The PCR product was purified from band cut out of the agarose gel using QIAquick gel extraction kit (Qiagen) and was directly sequenced using dye terminator chemistry on an ABI Prism 377 automated DNA sequencer with the PCR primers.

[View SNP found using this method](#)

[View SNP with population data obtained with this method](#)

FIG. 4E

SNP Publication Details

Submitter Handle: FGU-CBT

pmid:

MEDUID:

TITLE:

Single Nucleotide Polymorphism in SCA2 Gene.

AUTHOR:

CHOUDHRY, S.; BRAHMACHARI, S.K.

JOURNAL:

VOLUME:

SUPPL:

ISSUE:

I_SUPPL:

PAGES:

YEAR:

2000

STATUS:

1- unpulished

Searched PubMed by Author:

[CHOUDHRY, S.: BRAHMACHARI, S.K.](#)

[View SNP linked to this publication](#)

FIG. 4F



8/13

SNP Details

SNP:

Handle|local_snp_id: FGU-CBT | SKB.2K.1.2
NCBI Assay Id(ss #): 869705
Reference SNP Id(rs #): 696872

STS Information: Not submitted

From SNP Database:

Submitter Handle: FGU-CBT
Submitter Batch ID: SKB.2K.1
Release Data: Aug 2 2000 2:53PM
Molecular type: Genomic
No. of Chromosomes sampled: 215
Synonym defined:
Organism: Homo sapiens
Population: INDPOP
Submitter Method ID: SCA2-SNP
Citation:

Single Nucleotide Polymorphism in SCA2 Gene.

[View citation details](#)

NCBI Assay ID: 869705
Submitter SNP ID: SKB.2K.1.2
Synonyms:
LOCUSID: 6311
Submitter STS ID:
STS Accession: not available
GenBank Accession: U70323
Gene Name: Human ataxin-2 gene/Sphino cerebellar ataxia 2 (SCA2) gene
Length: 459

Flanking Sequence Information:

5' Assay: CTCCGCCTCA GACTGTTTTG GTAGCAACGG CAACGGCGGC GCGCGGTTTC GGCCCGGCTC
CCGGCGGGCTC CTTGGTCTCG GCGGGCCTCC CCGCCCCCTTC GTCGTCTGTC TTCTCCCCCT
CGCCAGCCCCG GCGCGCCCTC CGGCCGCGCC AACCCGCGCC TCCCCGCTCG GCGCCCC
(SEQ ID NO.: 17)
Observed: T/C
3' Assay: GCGTCCCCGC CGCGTTCCGG CGTCTCCTTG GCGCGCCCCG CTCCCGGCTG TCCCCGCCCCG
GCGTGCGAGC CGGTGTATGG GCCCCTCACC ATGTCGCT (SEQ ID NO.: 18)
3' Assay: GAAGCCCCAG CAGCAGCAGC AGCAGCAGCA GCAACAGCAG CAGCAGCAAC AGCAGCAGCA
GCAGCAGCAG CAGCCGCCGC CCGCGGCTGC CAATGTCCGC AAGCCCGGCG GCAGCGGCCT
TCTAGCGTCG CCCGCCGCC CGCCTTCGCC GTCCTCGTCC TCGGTCTCCT CGTCCTCGGC
CAC (SEQ ID NO.: 16)

Allele Frequency Information:

POP Batch Id: SKB.2K.1:
Submitter Handle: FGU-CBT
Submitter Method ID: SCA2-SMP
Citation:
Single Nucleotide Polymorphism in SCA2 Gene.
Handle| Population ID: FGU-CBT|INDPOP
No. of Chromosomes Sampled: 215
Allele: C = 0.293 / T = 0.707

FIG 5A



9/13

Reference SNP Record

NCBI SNP ID: rs695872

NCBI Resource Links

GenBank: U70323

LocusLink: no link established

Integrated Maps: *under construction*

Submitter records for this ID:

| Assay ID | Handle Local Submitter ID | Release Date |
|-----------------|-----------------------------|-------------------|
| <u>ss869705</u> | FGU-CBT SKB.2K.1.2 | Aug 2 2000 2:53PM |

Variation Summary:

Assay sample size (number of chromosomes): 430

Population data sample size (number of chromosomes):

Total number of populations with frequency data: 1

Total number of individuals with genotype data: 0

Average estimated heterozygosity: 0.414

Average Allele Frequency:

C: 0.293

T: 0.707

Validation Summary:

Marker displays Mendelian segregation: UNKNOWN

PCR results confirmed in multiple reactions: YES

Homozygotes detected in individual genotype data: UNKNOWN

Insufficient genotype data to compute the goodness of fit to Hardy-Weinberg

Insufficient data to compute individual x genotype consistency measures

Validation status: *under construction*

FIG. 5B



Inventor: Samir K. Brahmachari
Application No. 09/707,919
Title: METHOD OF DETECTION OF
ALLELIC VARIANTS OF SCA2 GENE
Client/Matter No. 39562-175772

10/13

SNP Publication Details

Submitter Handle:

pmid:

MEDUID:

TITLE:

Single Nucleotide Polymorphism in SCA2 Gene.

AUTHOR:

CHOUDHRY, S.; BRAHMACHARI, S. K.

JOURNAL:

VOLUME:

SUPPL:

ISSUE:

I_SUPPL:

PAGES:

YEAR:

2000

STATUS:

1- unpolished

Searched PubMed by author:

CHOUDHRY S.; BRAHMACHARI, S.K.

[View SNP linked to this publication](#)

FIG. 5C

SNP Population Details - 558

Submitter Population Handle: FCU-CBT

Submitter Population ID: INDPOP

Population Text:

Continent: Asia

Nation: India

[View SNP used on this population](#)

FIG 5D



11/13

SNP Method Details - 564

Submitter Method Handle: FGU-CBT
Submitter Method ID: SCA2-SNP
Method Text:

The region containing the SNP was PCR amplified using the primers SCA2-FP3 (5' CTCCGCTCAGACTGTTTTGGTAG 3') (SEQ ID NO: 1) and SCA2-RP3 (5' GTGGCCGAGGACGAGGAGAC 3') (SEQ ID NO. 2). Approximately 100ng of genomic DNA was amplified in a 50ml reaction volume containing a final concentration of 5mM Tris, 25mM KCl, 0.7mM MgCl₂, 0.05% gelatin, 20pmol of each primer and 0.5U of Taq DNA polymerase. Samples were denatured at 94oC for 3 min followed by 35 cycles of denaturation (94oC, 45sec), annealing (52oC, 30sec), extension (72oC, 45sec) and a final extension of 7 min at 72oC in a Perkin Elmer GeneAmp PCR System 9600. The PCR product was purified from band cut out of the agarose gel using QIAquick gel extraction kit (Qiagen) and was directly sequenced using dye terminator chemistry on an ABI Prism 377 automated DNA sequencer with the PCR primers.

View SNP found using this method

View SNP with population data obtained with this method

FIG. 5E

Submitter Contact Details

This batch's contact information:

handle: FGU-CBT
name: Shweta Choudhry
fax: +91-11-7257471
phone: +91-11-7416489
email: shwetachoudhry@hotmail.com
lab: Functional genomics Unit
institution: Centre for Biochemical Technology (CSIR)
address: Delhi University Campus, Mall Road, Delhi- 110007, India

Handle information for the lab:

handle: FGU-CBT
name: Prof. Samir K. Brahmachari
fax: +91-11-7257471
phone: +91-11-7416489
email: skb@cbt.res.in
lab: Functional genomics Unit
institution: Centre for Biochemical Technology (CSIR)
address: Delhi University Campus, Mall Road, Delhi- 110007, India

FIG 5F



Inventor: Samir K. Brahmachari
Application No. 09/707,919
Title: METHOD OF DETECTION OF
ALLELIC VARIANTS OF SCA2 GENE
Client/Matter No. 39562-175772

12/13

LOCUS HSU70323 4481 bp mRNA PRI 20-NOV-1996
DEFINITION Human ataxin-2 (SCA2) mRNA, complete cds.
ACCESSION U70323
VERSION U70323.1 GI:1679683
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Pulst, S.-M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X.-N.,
Lopes-Cendes, I., Pearlman, S., Starkman, S., Orozco-Diaz, G.,
Lunkes, A., DeJong, P., Rouleau, G.A., Auburger, G., Korenberg, J.R.,
Figueroa, C. and Sahba, S.
TITLE Moderate expansion of a normally biallelic trinucleotide repeat in
spinocerebellar ataxia type 2
JOURNAL Nat. Genet. 14 (3), 269-276 (1996)
MEDLINE 97051920
REFERENCE 2 (bases 1 to 4481)
AUTHORS Pulst, S.-M.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1996) Medicine, Cedars-Sinai, 8700 Beverly Blvd.,
Los Angeles, CA 90048, USA
FEATURES Location Qualifiers
source 1..4481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24.1"
gene 163..4101
/gene="SCA2"
CDS 163..4101
/gene="SCA2"
/standard_name="spinocerebellar ataxia type 2"
/codon_start=1
/product="ataxin-2"
/protein_id="AAV19200.1"
/db_xref="GI:1679684"
/translation="MRSAAAAPRSPAVATESRRFAAARWPGWRS LQRPARRSGRGGGG
AAPGPYPSPAAPPPPGPGPPSRQSSPPSASDCFGSNGNGGGAFRPGSRRLGLGGPPR
PFVVVLLPLASPGAPPAAPTRASPLGARASPPRSGVSLARPAPGCPRPACEPVYGPLT
MSLKPPQQQQQQQQQQQQQQQQQQQQPPFAANVRKPGGSGLLASPAAPSPSSSSV
SSSSATAPSSVVAATSGGGRPGLGRGRNSNKG L PQSTISFDGIYANMRMVHILTSVVG
SKCEVQVKNGGIYEGVFKTYS PKCDLVLDAAHEKSTESSGPKREEIMESILFKCSDF
VVVQFKDMDSSYAKRDAFTDS AISAKVNGEHKEKDLEPWDAGELTANEELEALENDVS
NGWDPNDMFRYNEENYGVVSTYDSSLSSYTVPLERDNSEEF LKREARANQLAEIESS
AQYKARVALENDDRSEEEKYTAVQRNSSEREHGSINTRENKYIPPGQRNREVISWGSG
RQNSPRMGQPGSGSMP SRSTSHTSDFNPNSGSDQRVVNGGVWPWSPCPSPPSRPPSR
QSGPNSLPRAATPTRPPSRPPSRPPSRPPSHPSAHGSPAPVSTMPKMSSEGPPRMSP
KAQRHPRNHRVSAGRGSISSGLEFVSHNPPSEAAATPPVARTSPSGGTWSSVVGVPRL
SPKTHRPRSPRQNSIGNTPSGPVLAS PQAGIIPTEAVAMP IPAASPTPASPNRAVT
PSSEAKDSRLQDQRQNSPAGNKENIKPNETSPSFSKAENKGISPVVSEHRKQIDDLKK
FKNDFRLQPSSTSESMDQLLNKNREGEKSRDLIKDIEPSAKDSFIENSSSNTSGSS
KPNSPSISPSILSNTEHKGREVTSSQGVQTSSPACKQEKDDKEKKDAAEQVRKSTLN
PNAKEFNPRSFSQPKPSTTPTSPRPQAQPSPSMVGHQQTTPVYTQPVCFAPNMMYPVP
VSPGVQPLYPIPMTPMPVNQAKTYRAVPNMPQQRQDQHHQSAMMHPASAAAGPPIAATP
PAYSTQYVAYSPQQFPNQPLVQHVPHYQSQHPHYVSPVIQGNARMMAPTHAQPGLV
SSATQYGAHEQTHAMYACPKLPYNKETS PSFYFAISTGSLAQYAHPNATLHPHTPH
QPSATPTGQQQSQHGGSHPA P SPVQHHQHQAALHLAS PQQSAIYHAGLAPTPPSM
TPASNTQSPQNSFPAAQQT VFTIHPSHVQPAYTNPPHMAHVPAHVQSGMVP SHPTAH
APMMLMTTQPPGGPQAALASALQPIPVSTTAHFPMTHPSVQAHHQQQL" (SEQ ID NO. 19)
BASE COUNT 1144 a 1380 c 1014 g 943 t

FIG 6A



13/13

ORIGIN

```
1 accccccgaga aagcaacccca gcgcgcgcgc cgctcctcac gtgtccctcc cggcccccggg
61 gccacctcac gttctgttc gctctgaccc ctccgacttc cggtaaagag tccctatccg
121 cacctccgct cccacccggc gcctcggcgc gcccgccctc cgatgcgctc agcggccgca
181 gctcctcgga gtcccgcggt ggccaccgag tctcgccgct tcgccgcagc caggtggccc
241 ggggtggcgt cgctccagcg gccggcgcg gggagcgggc gggggcggcg tggcgcgcc
301 ccgggaccgt atccctccgc cgcctctccc ccgcccggcc ccggccccc cccctcccg
361 cagagctcgc ctccctccgc ctccagactgt ttggtagca acggcaacgg cggcgcgcg
421 tttcggcccg gctcccggcg gctccttggt ctccgggggc ctccccgccc cttcgtcgtc
481 gtcccttctcc cctcgcgag cccggggccc cctccggccg cgccaacccg cgctccccg
541 ctccggcgccc gtgcgtcccc gccgcgttcc gggtctcctc tggcgcgccc ggtccccgg
601 tgtccccgccc cggcgtgcga gccggtgtat gggccctca ccatgtcgt gaagccccag
661 cagcagcagc agcagcagca gcaacagcag cagcagcaac agcagcagca gcagcagcag
721 cagccggcgcc ccgcggtgct caatgtccgc aagcccgcg gcagcggcct ctagcgtcg
781 cccgcgcgccc cgcttccgccc gctcctgccc gggagggccc gcctgggag aggtcgaaac
841 tccctcggtgg tcgcgcgac ctccggcgcc ggttttgatg gaatctatg aaatatgagg
901 agtaacaaag gactgcctca gtctacgatt tcttttgatg aagtacaagt gaaatatgga
961 atgggtcata tacttacatc agttgttggc tccaaatgtg gtgatttggg acttgatgcc
1021 ggtatatatg aaggagtttt taaaacttac agtccgaagt gtgaagaagt aatggagagt
1081 gcacatgaga aaagtacaga atccagttcg gggccgaaac gtacagttta ctagcgtcg
1141 attttgttca aatgttcaga ctttgttggc gtacagttta aagtgaatgg cgaacacaaa
1201 gcaaaaagag atgtttttac tgactctgct atcagtgcga aagtgaatgg actgaggtc
1261 gagaaggacc tggagccctg ggatgcaggt gaactcacag ccaatgagga actgaggtc
1321 ttggaaaaatg acgtatctaa tggatggggt cccaatgata tgtttcgata taatgaagaa
1381 aattatgggtg tagtgtctac gtatgatagc agtttatctt cgtatacagt gcccttagaa
1441 agagataaact cagaagaatt tttaaaacgg gaagcaaggg caaacagtt agcagaagaa
1501 attgagtcacaa gtgcccagta caaagctcga gtggccctgg aaaaatgata taggagttag
1561 gaagaaaaaat acacagcagt tcagagaaat tccagtgaac gtgaggggca cagcataaac
1621 actaggggaaa ataaatatat tccctctgga caaagaaata gagaagtcat atccctggga
1681 agtgggagac agaattccac gcgtatgggc cagcctggat cgggctccat gccatcaaga
1741 tccactttctc acacttcaga tttcaaccgg aattctgggt cagaccaaag agtagttaat
1801 ggaggtgttc cctggccatc gccttgcaca tctccttccc ctgcgccacc ttcgtcgtac
1861 cagtcagggtc ccaactctct tccacctcgg gcagccaccc ctacacggcc gccctccagg
1921 ccccccctgc ggccatccag accccctctc caacctctg ctcatggttc tccagctcct
1981 gtctctacta tgccataaac catgtcttca gaagggcctc caagtagtgc cccaaaggcc
2041 cagcgacatc ctcgaaatca cagagtttct gctgggaggg gttccatata cagtggccta
2101 gaatttgtat cccacaaccc acccagtga gctgggaggg gttccatata aaggaccagt
2161 cctcgggggg gaacgtgggt atcagtgggt gcagctactc ctccagtagc aaggaccagt
2221 catagaccca ggtctcccag acagaacagt attggaataa cccagattatc gccagttctt
2281 gcttctcccc aagctgggtat tattccaact gaagctggtg cccccagttg tccagctgca
2341 tctcctacgc ctgctagtcg tgcatcgaac agagctgtta ccccttctag tgaggctaaa
2401 gattccaggc ttcaagatca gaggcagaac tctcctgcag ggaataaaga aaatattaaa
2461 cccaatgaaa catcacctag cttctcaaaa aagaaattta agaattgatt taggttacag
2521 tctgaacata gaaaacagat tgatgattta ctactaaaca aaaatagaga gggagaaaaa
2581 ccaagttcta gttctgaatc tatggatcaa ccaagtgcga aggtattctt cattgaaat
2641 tcaagagatt tgatcaaaag caaaattgaa aagccgaata gccccagcat tccccttca
2701 agcagcagca actgtaccag tggcagcagc cctgaggtca ctcccaagg ggttcagact
2761 atacttagta acacggagca agagaaagac gataaggaag agaagaaaga cgcagctgag
2821 tccagcccag catgtaaaaca agatcccaat gcaaaggagt tcaacccacg ttccttctc
2881 caagttagga aatcaacatt cccaacttca cctcgccctc aagcacaccc tgcccatct
2941 cagccaaagc cttctactac aactccagtt tatactcagc ctgittgttt tgcaccaa
3001 atgggtgggtc atcaacagcc aagcccaggc gtgcaacctt tatacccaat acctatgacg
3061 atgatgtatc cagtcaccgt gagccacagt agagcagtag caaatatgcc ccaacagcgg
3121 cccatgccag tgaatcaagc caagcagcgt cagcagcggt cagcagcggt cccaccgatt
3181 caagaccagc atcatcagag tgcctatgat cagcagcggt acagtcctca gcagttccca
3241 gcagccaccc caccagctta ctccagcga tatgttgctc agcatcctca tgtctatagt
3301 aatcagcccc ttgttcagca ttgtccacat tctcagttc gcaccaccaa cacacgccc
3361 cctgtaatac agggtaatgc tagaatgatg gcaccaccaa cagcagcggt cagcagcggt
3421 gtatcttctt cagcaactca gtacgggggt catgagcaga ccttcttctt cagcagcggt
3481 cccaaattac catacaacaa ggagacaagc gctacccctg ccttcttctt accacatac
3541 tcccttgctc agcagtatgc gcaccctaac tggacagcag caaagccaac atgggtggaag
3601 cagccttcag ctacccccac tgcagaccag gccgcccagg cctcctcatc ctccaccctc
3661 cccagtcctg ttcagcaca ccacgcgggg cttgcgccc cctccagcag cacaacagac
3721 cagcagcagt cagccattta ccacgcgggg ttcccagcag cccacatggt cccacatggt
3781 gcctccaaca cgcagtcgccc acagaatagt accaaccac ccacatggt cccacatggt
3841 atccatcctt ctacggttca gccggcgctt tctcatccaa ctgcccagtc cgcctcaaa
3901 caggctcatg tacagtcagg aaggttccct cagggcgccc tcgctcaaa
3961 ctaatgacga cacagccacc cggcggtccc ccttatatga cgcacccttc aaaggccaaa
4021 cccattccag tctcgacaac agcgcatttc ggaggaaccg aaaggccaaa aatttccatt
4081 caccaccaac agcagttgta aggtgcctt ggaggaaccg aaaggccaaa aatttccatt
4141 ccttctactg cttctaccaa ctggaagcac agaaaactag ccaataagaa tgctaacagt
4201 taaaatatat atgttgattt ctgttaacat taggaacttg ggggctattc cataattcca
4261 ttggaacgag ttggaacgag tagagcatt ggggctattc gccgaaactg gaagttattt
4321 tatgctgttt cagagtcctc aggtaccacc agctctgctt acatcagcta gcaaaagaa
4381 attttttaac aaccttgaa ctcatgaac aaaaaaaa a a (SEQ ID NO: 20)
4441 gattcttgct gctattactg ctaaaaaaa aaaaaaaa a
```

FIG 6B